Price-Sensitive Application Adaptation in Deadline-Based Networks*

Xiao Huan Liu
Department of Computer Science
University of Manitoba
Winnipeg, MB R3T 2N2,Canada
liuxh@cs.umanitoba.ca

Yanni Ellen Liu
Department of Computer Science
University of Manitoba
Winnipeg, MB R3T 2N2,Canada
yliu@cs.umanitoba.ca

ABSTRACT

In deadline-based networks, the delay performance observed by real-time data largely depends on the traffic deadline and the level of load along the data path. To prevent greedy users from gaining an advantage by specifying arbitrarily urgent deadlines and to aid in network load control, a delay pricing and charging scheme was developed for real-time delivery in deadline-based networks. Using this scheme, users experiencing different delay performance are charged differently. In respond to an earlier charge, a price-sensitive adaptive application with a limited budget constraint may adjust its traffic accordingly by varying its traffic deadline requirements or traffic load intensity. In this poster, we show through simulation that the developed pricing and charging scheme easily enables such adaptations, which in turn may significantly improve the performance of real-time delivery in deadline-based networks.

1. INTRODUCTION

Deadline-based network resource management [4] is a framework that has been developed to support real-time data delivery in packet-switched networks. In this framework, each application data unit (ADU) is given a delivery deadline by the sending application. It represents the time at which the ADU should be delivered at the receiver. The ADU deadlines are mapped to packet deadlines, which are carried by packets and used by routers for channel scheduling. Deadline-based channel scheduling algorithm is used inside routers. It was shown that deadline-based scheduling achieves superior performance to FCFS (First-Come First-Served) with respect to the percentage of ADUs that are delivered on time [2].

In deadline-based networks, the delay performance observed by real-time data largely depends on the deadline that it carries. If one is free to specify the ADU deadline, a sender may try to gain an advantage by using arbitrarily tight deadlines. Besides deadline urgency, the delay performance in deadline-based networks is also affected by the load conditions along the ADU path. When the load is light, the delay performance is good. When the load is heavy, congestion may occur; the delay performance deteriorates. To prevent greedy users from specifying arbitrarily urgent deadlines and to aid in network load control, a novel delay pricing and charging scheme has been developed [1].

In this scheme, a market-based approach from the field of microeconomics was taken. Each channel periodically computes a channel *delay price* based on the relation between the delay supply and the delay demand. The delay supply S is derived from the total time it takes to service all the packets in a *price update interval*. The delay demand D is derived from the deadline information carried by these packets. At the end of the update interval n, the channel delay price p for the update interval $n+1$ is defined as: $p_{n+1} = \{p_n + \sigma * (D - S)/S, 0\}^+$, where σ is an adjustment factor that can be used to trigger faster or slower responses of the channel price to the amount $D - S$. The end result of this scheme is that the channel price is higher (i) when the deadline urgency is higher, and (ii) when the load is heavier.

Given this pricing scheme, packets and ADUs are charged as follows. At each channel, upon each packet departure, the packet response time (which is defined as the sum of the queueing delay, the packet transmission time, and the channel propagation delay) at this channel d_a is calculated. Let p be the current channel delay price. The packet charge g at this channel is defined as: $g = p/d_a$. A new packet header field called "accumulated charge" is defined to keep track of the total delay charge incurred by this packet at all channels along its path. If a packet arrives at the receiver on-time, the value of this field is retrieved and is taken as the packet charge. If an ADU is delivered on-time, its ADU charge is the sum of all its packets' charges. Late packets and ADUs are not charged.

In respond to the charges assigned by the network to earlier transmissions, a price-sensitive adaptive application with a limited budget constraint may adjust its traffic by varying its traffic deadline requirements or load intensity. In this poster, we show through simulation that the developed pricing and charging scheme easily enables such adaptations, which in turn may significantly improve the performance of real-time delivery in deadline-based networks.

2. PRICE-SENSITIVE ADAPTATION

Our goal is to demonstrate that the aforementioned pricing and charging scheme easily enables application adapta-

[1] Refer to http://www.cs.umanitoba.ca/∼liuxh for more information.

tions, which in turn can lead to much improved performance. In what follows, we focus on those applications that may adapt their traffic requirements. We assume that application adaptation is performed periodically and the adaptation intervals coincide with the price update intervals. We assume that each user has a fixed amount of budget for every price update interval, the ADU charges are made available to the senders via application layer acknowledgments, and each user records the total charge that is received in every price update interval.

Based on the charge in the previous interval, one or more of the following three traffic attributes can be adjusted: the average ADU deadline, the average ADU inter-arrival time, and the average ADU size. When the budget is lower than the charge received in the previous interval, a user may lower the traffic requirements in terms of less urgent deadlines, a lower ADU arrival rate, and smaller ADUs. When deadlines are less urgent, the response time would increase, which results in a lower charge. When the ADU arrival rate is lowered, less ADUs are sent, thus less ADUs are charged. When ADUs are smaller, the load level can be reduced, thus the amount $D - S$ would decrease, which results in lower prices and lower charges. Conversely, when the budget is greater than the charge received in the previous interval, a user can raise the requirements of his/her traffic.

3. PERFORMANCE EVALUATION

We use discrete event simulation to evaluate the network performance when with and without application adaptation. A 6-node network model is used. The capacity of each channel is assumed to be 45 Mbit/sec. The propagation delay on each channel is 0.01 sec. Fixed routing is used. The total buffer size on each channel is 1MByte. For simplicity, we ignore the processing delay inside routers and end-systems.

Denote each source/destination pair as a traffic class. There is one session per traffic class, each session continuously generates ADUs during the simulation. Each session initially sets up three attributes: 1) end-to-end deadline E, 2) mean ADU interarrival time I, and 3) ADU size parameter θ. All sessions' initial E is assumed to be 80 ms. The ADU interarrival times are exponentially distributed with mean I. All sessions' initial I is 0.2 second. This corresponds to 88% utilization on the bottleneck when without user adaptation. The ADU size is the product of θ and a random number z, in bytes. z is generated from two ranges with equal probability: Uniform(500, 1500), and Uniform(1500, 500000).

In our experiments, we assume that a session adapts the three attributes as follows. Define an adaptation parameter α. Let the value of an attribute in the update interval n be v_n. The value of the attribute in the update interval $n+1$, v_{n+1} will be calculated by either $v_{n+1} = v_n * (1 + \alpha)$ or $v_{n+1} = v_n * (1 - \alpha)$. In our experiments, an α value of 0.2 is used. The price update interval is assumed to be 3 seconds' long. All sessions' budget per interval is assumed to be 1. Note that we do not associate the charge with any concrete monetary value and leave this choice to network operators.

Four experiments are performed: one without adaptation, in the other three, the three attributes: E, I, and θ are adapted respectively. The performance measures used are the aggregated ADU on-time rate and the ADU on-time throughput (in number of ADUs per second). The results from our simulation are shown in Table 1. It can be observed that the network performance can be significantly improved after application adaptation.

Table 1: Performance without and with adaptation

Adapting attributes	on-time rate	throughput
Without adaptation	0.6104	227.15
Adapting deadline E	0.6493	241.63
Adapting inter_arr_t I	0.7249	241.1
Adapting ADU size θ	0.8432	313.78

Our simulation results also show that the channel delay price on the bottleneck can be well controlled when with user adaptation, while the price keeps increasing if without adaptation. This can be observed from the bottleneck price dynamics plot in Figure 1.

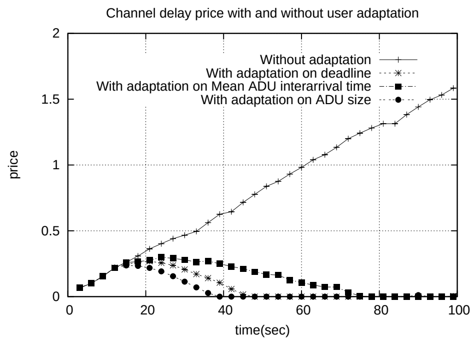

Figure 1: Prices when without and with adaptation

4. RELATED WORK

Price-sensitive application adaptation has been studied previously, see e.g., [3]. However, differ from existing studies where pricing and adaptation of *bandwidth* is concerned, we focus on traffic *delay* pricing and adaptation. This is made available by the deadline-based framework in which each packet carries its delay requirement.

5. CONTRIBUTIONS

We study the effect of application adaptation that is enabled by a novel delay pricing and charging scheme in deadline-based networks. We assume that price-sensitive users with limited budget constraints will adapt their traffic requirements based on the charges received earlier. We show through simulation that the delay performance of the entire network can be improved as the result of such adaptations.

6. REFERENCES

[1] Y. E. Liu and X. H. Liu. A Delay Pricing Scheme for Real-Time Delivery in Deadline-Based Networks. In *Proc. of WINE*, 2005. To appear.

[2] Y. E. Liu and J. W. Wong. Deadline based channel scheduling. In *Proc. of the IEEE Globecom'2001*, pages 2358–2362, San Antonio, Texas, Nov. 2001.

[3] X. Wang and H. Schulzrinne. Performance study of congestion price based adaptive service. In *Proc. of NOSSDAV*, pages 1–10, Chapel Hill, NC., Jun. 2000.

[4] J. W. Wong and Y. E. Liu. Deadline based network resource management. In *Proc. of ICCCN'2000*, pages 264–268, 2000.